

Patent Application

for

HEAP MEMORY MANAGEMENT

Inventor(s):

JACK M. BAYT

Prepared By:

Jason S. Feldmar
Gates & Cooper LLP
Howard Hughes Center
Suite 1050
6701 Center Drive West
Los Angeles, California 90045

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit under 35 U.S.C. Section 119(e) of the following co-pending and commonly-assigned U.S. provisional patent application(s), which is/are incorporated by reference herein:

- 5 [0002] Provisional Application Serial No. 60/455,899, filed March 19, 2003, by Jack M. Bayt, entitled "HEAP MANAGEMENT" attorneys' docket number 30566.297-US-P1.

BACKGROUND OF THE INVENTION

- 10 1. Field of the Invention.

[0003] The present invention relates generally to managing memory, and in particular, to a method, apparatus, and article of manufacture for using a heap to manage memory.

- 15 2. Description of the Related Art.

[0004] Files and file systems provide mechanisms for long term storage of an application and are supported by most operating systems. Programming languages usually provide a layer application programming interface (API) to use to access files that then internally uses the underlying operating file system. Generally, in the prior
20 art, a file and file systems are linear in nature (i.e., they are accessed in a linear manner). Thus, the application is able to write data to a point (or read data from a point) in a sequential manner. However, file systems may also be randomly accessed which allows one to "seek" to a point but from that point, access is sequential in

nature. For example, a file pointer may be randomly set to point to the location where data is either written to or read from in a linear/sequential manner. Accordingly, while the file systems are linear in access, the file pointer may be positioned randomly.

5 **[0005]** Once a file pointer is set, an application is able to read or write data in a linear manner. However, the application is responsible for keeping track of what data is stored at what file pointer position. Further, when data is inserted into the linear file, the entire file after the insertion point must be rewritten. Accordingly, the prior art requires applications to maintain significant knowledge about file systems while
10 inserting and removing data inefficiently.

[0006] FIG. 1 illustrates the use of a prior art linear file access system. The file system allows a user to seek 102 to a particular memory offset location N (thereby providing the ability to randomly position a pointer). Thereafter, the file is linear in nature stored at bytes N 104, N+1 106, and N+2 108.

15 **[0007]** The physical file system used by many operating systems is often a collection of blocks with linkage information between the blocks that is controlled by a linear access array. FIG. 2 illustrates a prior art file system controlled by such a linear access array. For a given offset in the file, the index array of pointers 202 is used to point to the block 204-208 that contains the given data. The requested data
20 may then be accessed in the block.

[0008] When storing and retrieving files, memory is consumed. Various methods have been developed for managing memory. For example, some prior art applications may utilize heaps to manage memory. A heap is a term used to describe a pool of

memory available to an application. Applications request a block of memory from the heap for use (allocation). The application may later release the memory back to the heap (deallocation). The implementation of a heap system is generally a requirement for all operating systems.

5 [0009] A common feature of most heaps is that they keep track of all blocks that were previously allocated and then later deallocated (so they may be reused by later allocation requests). Such a list of deallocated blocks is commonly called a free-list. When an allocation request is made, the heap will attempt to reuse a deallocated block from the free-list prior to requesting new memory from the operating system. In this
10 regard, the free-list is searched for a block that satisfies the allocation request. In the prior art, the free list is searched in a linear manner.

[0010] In one prior art example, a heap is broken up into multiple chunks – each representing approximately a megabyte of free memory. In each chunk, the free blocks are stored sequentially. Thus, searching for a particular size or finding the
15 smallest block that will satisfy the desired allocation request is of order N – linear.

[0011] The prior art may also utilize a heap configured as a binary tree where each node in the tree represents a block of memory of a certain size. Two links exist for each node. One link points to memory blocks smaller in size while a second link may point to memory blocks of equal or greater size. The problem with such binary free-
20 list trees is that applications tend to have many objects of the same size in the free list. Accordingly, when traversing a binary tree to find a satisfactory block of memory, search times may be near the same as a linear search.

[0012] In addition to the above, it is common for programmers to have a coding

error that results in a heap being asked to deallocate a block that the heap never allocated. Heaps generally do not detect this kind of error and thus they cause an application to crash. Sometimes, there are different heaps used – one used during development that has significant overhead in integrity checking, and another used for the end product that has no checking. While such a solution may help reduce errors, in many cases, there are significant differences in runtime behavior of an application between development mode and end product mode. Accordingly, what is needed is an ability to quickly and efficiently perform integrity checking at runtime that does not utilize significant overhead.

10

SUMMARY OF THE INVENTION

[0013] Heaps are a term used to describe a method to allocate and manage memory for use by an application. Applications request a block of memory from the heap for use (allocation). Later, the application releases the memory back to the heap (deallocation). Usually, a heap also maintains a free list of those blocks of memory that have been deallocated so they may be reused by later allocation requests.

15

[0014] File systems are a way to manage long term storage for an application. Prior art file systems utilize a file pointer that points to the location where data is either written to or read from in a linear manner. File systems are linear in access but the file pointer may be positioned randomly.

20

[0015] One or more embodiments of the invention combines the concepts of a heap and a file system. Instead of a file pointer, an application requests from the file heap an object of a desired size and then is able to read/write data to that object in a

random manner. A file may be broken up into blocks and a tree organization is used to manage the blocks. To insert data into a file, the block is merely broken at the insertion point, and the new block is inserted as a node in the tree. Random access may be simulated by mapping the linear address of a file through the tree structure to
5 map to the actual node used for that block.

[0016] In addition, the invention provides for the use of a tri-linked free list. Instead of a linear search, a binary tree is used for the search wherein nodes of the same size are stored outside of the tree. Accordingly, the tri-linked tree only contains a single reference to a particular block size in the free list.

10

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Referring now to the drawings in which like reference numbers represent corresponding parts throughout:

[0018] FIG. 1 illustrates the use of a prior art linear file access system;

15 [0019] FIG. 2 illustrates a prior art file system controlled by a linear access array;

[0020] FIG. 3 is an exemplary hardware and software environment used to implement one or more embodiments of the invention;

[0021] FIG. 4 illustrates the organizational structure for a file system modeled as a heap in accordance with one or more embodiments of the invention;

20 [0022] FIG. 5 is a flow chart illustrating the use of a heap as a file system in accordance with one or more embodiments of the invention;

[0023] FIG. 6 illustrates a tri-linked free list in accordance with one or more embodiments of the invention; and

[0024] FIG. 7 is a flow chart illustrating the traversal of a tri-linked tree in accordance with one or more embodiments of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0025] In the following description, reference is made to the accompanying drawings which form a part hereof, and which is shown, by way of illustration, several embodiments of the present invention. It is understood that other

5 embodiments may be utilized and structural changes may be made without departing from the scope of the present invention.

Hardware Environment

[0026] FIG. 3 is an exemplary hardware and software environment used to
10 implement one or more embodiments of the invention. Embodiments of the invention are typically implemented using a computer 300, which generally includes, inter alia, a display device 302, data storage devices 304, cursor control devices 306, and other devices. Those skilled in the art will recognize that any combination of the above components, or any number of different components, peripherals, and other devices,
15 may be used with the computer 300.

[0027] One or more embodiments of the invention are implemented by an operating system or memory manager program 308. Generally, the memory manager 308 comprises logic and/or data embodied in or readable from a device, media, carrier, or signal, e.g., one or more fixed and/or removable data storage devices 304 connected
20 directly or indirectly to the computer 300, one or more remote devices coupled to the computer 300 via a data communications device, etc. Further, the memory manager 308 controls the file system and manages the memory for other applications executing on computer 300.

[0028] Those skilled in the art will recognize that the exemplary environment illustrated in FIG. 3 is not intended to limit the present invention. Indeed, those skilled in the art will recognize that other alternative environments may be used without departing from the scope of the present invention.

5

File System Modeled as a Heap

[0029] As described above, with a prior art file system, significant overhead is required to maintain a file system (including knowledge regarding memory offsets). Examples of such prior art file systems include NTFS (NT File System) and/or the
10 FAT (file allocation tables) system. In addition, it is difficult to insert data into the middle of the file without significant overhead. In order to insert data into the middle of a file, you must move all data after the insertion point to a new location in the file to make room for the data to be inserted.

[0030] One or more embodiments of the invention combine the ideas of a heap and
15 a file system. The invention maintains a heap that represents the file system. Utilizing a heap in this manner enables an efficient mapping and separation between the logical address used by a file and the physical storage location.

[0031] FIG. 4 illustrates the organizational structure for a file system modeled as a heap in accordance with one or more embodiments of the invention. The bottom
20 portion of FIG. 4 illustrates the heap 400 managing the memory blocks 402-406. Instead of a file pointer, an application 402 may request from the file heap 400, an object of a desired size. The application 402 may then read/write data to that object in a random manner. There is no file pointer since each object is unique and

independent. File heap blocks 402-406 may be allocated and deallocated in a random manner, and the file heap 400 maintains all information related to actual location without any knowledge of the application 402. The file heap 400 may physically store the data in a traditional file or in any other manner allowed by the hardware operating system.

[0032] Accordingly, instead of an application 402 maintaining pointer and/or offset information, the application 402 merely retrieves an object from the heap 400. The heap 400 therefore manages all of the memory blocks. In this regard, the heap 400 acts as an application programming interface (API) to the application 402. The application 402 may insert data into a file, delete data from a file and perform any desired memory operation without knowledge of the underlying memory or file system status. Accordingly, the file being accessed looks like a heap and does not appear as a linear file for which offset and memory storage must be maintained.

[0033] With a file system modeled as a heap 400, the issue arises as to how legacy applications 410 may utilize the new file system. To reprogram existing legacy programs to directly interface with a heap 400 instead of maintaining file offsets and managing the memory may take considerable time and effort. Accordingly, it would be beneficial to allow legacy applications (with minor or no changes) to utilize the new heap-based system while permitting newer applications 402 to interact directly with the heap 400.

[0034] One or more embodiments of the invention provide the ability for legacy applications to utilize a heap-based file system. In legacy applications, a file may be accessed in a linear manner. However, the underlying file system may no longer be

linear. To allow the application 410 to continue to use linear addressing, the invention provides a file address mapping tree 412 that translates linear file addresses into heap block references. Random access may be simulated by mapping the linear address of a file through the tree structure to map to the actual node used for that
5 block (i.e., the file address mapping tree 412).

[0035] Accordingly, when an application 410 requests memory at a particular linear memory address, the file address mapping tree 412 converts the request to a heap block reference which is passed onto the heap 400 for further processing. Thus, while the legacy application 410 appears to be working with a linear memory block, the
10 heap 400 is actually managing the memory. In this regard, by having a heap represent the file system, there is an efficient mapping and separation between the logical addresses used by an application 410 and the physical storage location (managed by the heap 400).

[0036] Utilizing the heap 400 to represent the file system, and by having an address
15 mapping layer 412 above the heap 400, the insertion and deletion of address spaces and data in a file are more easily facilitated. Since the file is broken up into blocks, and a heap tree organization 400 is used to manage the blocks, one may insert data into a file simply by breaking the block at the insertion point, and inserting the new inserted data as a node in the heap tree 400.

20 [0037] Thus, to insert data randomly into a file, an existing block 404-408 is merely broken up, the heap tree 400 is updated, and the mapping layer 412 is adjusted.

Accordingly, space may be inserted into the file by splitting the file heap 400 at the appropriate place, inserting a new block into the heap 400, at most copying a partial

block to a new heap node at the insertion point, and adjusting the file address mapping 412 accordingly.

[0038] Areas of a file can be quickly deleted merely by adjusting the file address mapping tree 412 and deleting the associated blocks 404-408 from the heap or
5 reducing the size of partial blocks as needed.

[0039] Utilizing the heap 400 to represent a file system also provides additional advantages. For example, one may compress or encrypt (or both) 414 the logical user data prior to saving to the heap 400. Prior art file systems implement compression and encryption by compressing or encrypting an entire file. To access an encrypted
10 file, the entire file is first completely uncompressed or unencrypted and saved to a temporary file. Operations are then performed on the temporary file. When the file is closed, the entire file is recompressed or reencrypted.

[0040] By using the heap 400, only the block(s) that contain the physical data that is modified need to be recompressed or reencrypted 414. Accordingly, each block may
15 be independently compressed and decompressed (rather than the entire file).

Modified data most likely compresses or encrypts to a new size. In response, the heap 400 may allocate or reuse a block of the appropriate size for the modified data without reprocessing the entire file.

[0041] FIG. 5 is a flow chart illustrating the use of a heap as a file system in
20 accordance with one or more embodiments of the invention. At step 500, a file is broken up into two or more memory blocks. At step 502, the memory blocks are managed as nodes in a heap tree 400. Each node in the tree has a heap block reference. At step 504, a request to access memory (e.g., to insert or delete data) at a

linear file address is received. At step 506, the linear file address is translated into an appropriate heap block reference to access the memory block.

[0042] As described above, the translation may utilize a file address mapping tree 412 that maps linear addresses to heap block references. Further, when a block is inserted or deleted from the heap tree 400, the file address mapping tree 412 is updated. When deleting data from a file, the file address mapping tree 412 is updated by deleting an associated block from the heap 400, reducing a size of partial blocks as needed, and adjusting the file address mapping tree 412 accordingly. When inserting data into the heap 400, the memory blocks may be broken at the insertion point, and new data is inserted as a node in the heap tree 400.

Tri-Linked Free List for a Heap

[0043] As described above, heaps may be required to maintain a list of all allocation units that are no longer in use (i.e., deallocated memory units). This list is referred to as a “free list”. When a heap 400 is asked to allocate a block, the heap 400 searches the free list first to reuse existing unused blocks before expanding by allocating a new block. Prior art heaps 400 may use either free lists or a linked linear list of free blocks, or a combination of the two. Further, prior art free lists may be maintained in the form of a binary tree. The problem with using a binary tree for a free list is that applications tend to have many objects of the same size in the free list. Such repetitiveness causes binary trees to have search times near the same as a linear search.

[0044] One or more embodiments of the invention utilize a tri-linked free list. A

tri-linked free list is a tree that points to blocks smaller than itself, larger than itself, and the same size as itself. In this regard, the tri-linked tree may also be viewed as a binary tree, wherein nodes of the same size are stored outside of the binary tree. Such external storage provide the ability for each node in a binary tree to only contain a
5 single reference to a particular block size in the free list.

[0045] FIG. 6 illustrates a tri-linked free list in accordance with one or more embodiments of the invention. Each node 602-608 in the tri-linked free list 600 represents a block of memory of a particular size. Three links may be used in the tri-linked free list 600. The first link 602 points to blocks smaller than a current block
10 size N 602. A second link 606 points to blocks equal to the current block size N 602. The third link 608 points to blocks larger than the current block size N 602. Accordingly, in contrast to the prior art, the tree does not need to search all blocks of equal size 606. Instead, the blocks of the same size 606 (which are easily and efficiently linked together) are eliminated from the search. Since allocation sizes tend
15 to repeat, elimination of these equal blocks may dramatically increase search and edit performance.

[0046] FIG. 7 is a flow chart illustrating the traversal of a tri-linked tree in accordance with one or more embodiments of the invention. The process illustrated in FIG. 7 determines the memory unit that best fits the request. Alternative
20 embodiments may follow the first fit or other analysis for satisfying the memory requested.

[0047] At step 702, a request is received for memory. The request indicates that size of the memory block needed. The tree traversal begins with a first/current node.

A determination is made at step 704 regarding whether the current node being examined is large enough to satisfy the request.

[0048] If the request is large enough, the reference to the current block (i.e., the link to memory blocks equal to the current block size) is stored at step 706. Such a storage may overwrite any reference already stored at that location. The traversal then continues with a determination of whether there are any nodes left at step 708. If there are nodes in the tri-linked free list left, the child link to the memory units smaller than the current block size is followed at step 710 and the process returns to step 704. However, if there are no nodes left in the tree (as determined at step 708), the process is complete at step 712 and the stored link is used to satisfy the memory request (i.e., the memory block pointed to is allocated to the requesting application). Alternatively, if no reference is stored, a new memory block may be allocated.

[0049] If the current node is not large enough to satisfy the request (as determined at step 704), a determination is made as to whether the tree has been completely traversed at step 714. If there are no nodes left, the process is complete at step 712 and the stored reference to the appropriate block size is allocated. Alternatively, if there is no reference stored, a new memory block may be allocated/requested from the operating system. If there are nodes left as determined at step 714, the tree traversal continues by advancing to the child node with a block size larger than the current node at step 716 and returning to step 704.

[0050] As described above, the entire tree 600 is traversed to find the smallest memory unit/block that satisfies the request. Further, unlike the prior art, the process does not examine or search nodes with memory units of equal size (since they are all

linked together). Accordingly, such a repetition is avoided/eliminated.

Heap Controlled Blocks

5 [0051] As described above, the prior art heaps may encounter problems when being asked to deallocate a block that the heap never allocated. Accordingly, a common function of a heap is to be able to easily and efficiently identify blocks that the heap controls. In the prior art, a special signature at the beginning (or end) of a block may be used and compared to a list of blocks controlled by the heap.

10 [0052] One or more embodiments of the invention provides for a heap that maintains a bitmap/bitmask where every bit identifies a range of addresses that are owned by the heap. Prior to any deallocation request, the heap is able to perform a simple shift operation on the address and use the resulting value as an index into the bitmap to see if the address range is owned by the heap.

15 [0053] An example that may utilize this invention is a heap in a Windows™ Win32 environment. Windows™ allocates blocks on 64k byte boundaries. Thus, the heap may shift the address of any pointer to the right 16 bits to obtain a block index number:

$$\text{Block Index} = \text{Block Pointer} \gg 16.$$

20 The resulting index is used into an array of bits (bitmap) to quickly see if the block is owned by the heap or not. A bit is set if the block is used by the heap, and is clear if it is not used.

[0054] A shift operation to perform heap ownership of memory blocks is efficient and may be fast enough to use in production builds of a program.

Conclusion

[0055] This concludes the description of the preferred embodiment of the invention.

The following describes some alternative embodiments for accomplishing the present
5 invention. For example, any type of computer, such as a mainframe, minicomputer,
or personal computer, or computer configuration, such as a timesharing mainframe,
local area network, or standalone personal computer, could be used with the present
invention.

[0056] The foregoing description of the preferred embodiment of the invention has
10 been presented for the purposes of illustration and description. It is not intended to be
exhaustive or to limit the invention to the precise form disclosed. Many
modifications and variations are possible in light of the above teaching. It is intended
that the scope of the invention be limited not by this detailed description, but rather by
the claims appended hereto.